DOCUMENT RESUME

ED 040 991 TE 001 880

AUTHOR Cavanagh, Gray C.; Trip, Gus Van Vierssen
TITLE Evaluation and the New Freedom in Grade XIII English.
INSTITUTION Ontario Council of Teachers of English, Toronto (Ontario).
PUB DATE 70
NOTE 8p.
JOURNAL CIT English Exchange; v12 n2 p22-23, 26-30, 32 W 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$0.50
DESCRIPTORS Academic Freedom, Achievement Tests, College Admission, College Entrance Examinations, Educational Objectives, *English Instruction, Grading, High School Students, Individualized Programs, *Individual Needs, Individual Tests, Inservice Teacher Education, Objective Tests, *Standardized Tests, *Student Evaluation, Student Needs, *Testing, Testing Problems, Test Reliability

ABSTRACT

Tests are inevitably designed to measure students and therefore will affect both teaching techniques and the student's concept of what is important in learning. Standardized, objective high school English tests, which neglect the individual's needs or accomplishments, limit classroom freedom and the natural course of the curriculum. An integral component of the teaching process meaningful evaluation must grow out of the teaching situation and relate to such learning objectives as personal sensitivity, imagination, and creativity. An evaluation profile, whether designed to facilitate grading or college entrance decisions, is ideally based on multiple criteria (i.e., creative writing, individual projects, and participation in debates), accumulated by more than one method (i.e., multiple marking--weighing the reactions of several assessors to student work), and carried out over a long period of time. (MF)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED040991

Evaluation and the New Freedom in Grade XIII English

GRAY C. CAVANAGH *and* GUS VAN VIERSSEN TRIP

In our educational system, testing tends to be considered a self-contained process designed solely for the purpose of evaluating the student and assigning him an appropriate mark. The act of reducing the student's achievement to a number leads to a divorce in the teacher's mind—of the examination and its purpose from the purpose of teaching English. The formal examination with its routine of classroom silence, time limit, strict rules, dispensation of regular school timetable, and emphasis upon detailed recall of the course studied serves to strengthen this impression of special separation.

Testing needs to be considered an integral component of the teaching process—not a divorcee capable of segregation from the holy state of puristic methodology. The reasons may be classified:

1. A test designed to measure cannot escape becoming a teaching technique.
2. A test designed to reveal to the teacher certain "facts" about the student inevitably reveals something to the student and not infrequently the wrong "facts".
3. Tests designed simply to meet the practical need for report card marks also shape the student's concept of what is important in the learning process.
4. Uniform tests imposed across a school system solely to secure identical standards in all schools nevertheless directly control curriculum content in individual schools.
5. The type of examination imposed upon teachers across a school system to secure uniformity of standards appreciably dictates the individual teacher's methodology and emphasis.
6. The imposition of an externally-set examination, objective-answer or otherwise, to classify large numbers of students for university entrance inevitably eliminates the freedom and ability of the indi-

TE001 880

vidual school, English staff or teacher to design a course and methodology suited to the students' needs, interests and talents.

Ideally, evaluation in English should not shape methodology or curriculum; on the contrary, evaluation should grow out of the teaching situation, should serve as an integrated component in the teacher's battery of instructional methods and relate naturally to the teaching purposes structured in the course of study. The curriculum and teaching approach should in turn be determined by the teacher's concept of teaching purpose and assessment of the needs, interests and talents of his students.

The 1967-68 trend in testing for Grade XIII was exactly that—Department of Education Memorandum 98 placed all Grade XIII secondary school teachers in the position where they would have authority to determine the purpose and method of evaluation to be used in the classroom. Memorandum 98 appeared on the 18th of May, 1967, and on the 5th of December, 1967, Memorandum 28 appeared to reinforce the new school freedom and responsibility, with this statement:

"Departmental examinations have been discontinued in order to strengthen and improve education by permitting the teacher to consider the educational needs of each student as an individual . . . the practice of holding uniform examinations for more than one school is not recommended since it would in effect continue the disadvantages inherent in any external examination."

In the 1967 A.S.C.D. Yearbook, in an article entitled "The Evaluation We have", F. S. Bedell interprets the implications of the new policy:

"The abolition of the written (external) examination will release, in a large proportion of pupils, creative abilities which have been smothered atrophied, or consciously suppressed by the doctrine that schooling and the acquisition of information are synonymous and have the same boundaries."

At the O.S.S.T.F. Grade XIII Conference on October 20-21, Archie T. Carnahan, Group Chairman of the Social Sciences, said in the keynote address:

"*Evaluation* . . . is much broader than testing. It must encompass every objective valued by the school. It includes much that depends on personal sensitivity and intuition . . ."

The concept of the teacher as the most reliable arbiter of the student's progress is the key point in this remark. Robert D. Tarleck in the January '67 issue of the *ATA Magazine* states:

"I know of no study to date which does not indicate that distributive evaluation furnishes a more precise measurement of learning than that based on one testing day."

An evaluation profile based on multiple criteria, different methods of assessment and more than one short period of time is the keynote of this remark.

The answer to the question of how to evaluate is specifically this:

1. In the teaching of English, there should never be a single criterion or sample of work by which the student is measured. Two educational researchers, Vernon and Millican, demonstrated this fact in 1954 when they investigated the consistency of seven pieces of writing by college students. They concluded that "a consistent English ability, recognisable by different examiners from different small samples of students' work, can barely be said to exist", and thus they threw "very grave doubt on the common practice . . . of trying to assess English ability in general from a single essay marked by a single examiner." The year's standing in English, therefore, should not be determined by any June examination alone or by an objective answer test.
2. All students should be evaluated on the principle of total performance; that is, the day by day performance considered over the entire school year. The application of this and the previous principle means that a student's final mark might be some synthesis of 30 different units of work—oral presentations, expositions, creative writing, individual projects, research essays, brief analysis of literature studied, evidence of initiative, participation in debates, panel discussions, and other combinations of the student's contribution to the learning process.
3. Methods of evaluation should be designed to leave the student with an awareness of what is truly important in the learning situation. Facts should not be stressed, or rote memorization will be assumed important to success. In contrast, if the ability to marshal facts as evidence and to use them to support a reasoned explanation is emphasized, the student will come to value the abilities to formulate a view and support it convincingly. Likewise, if individual interpretation and vital expression are emphasized, the student will cultivate initiative, thinking ability and skill in writing.
4. To some extent and on certain occasions the students should be involved in the evaluation process. Through self-evaluation and the evaluation of the work of classmates, the student consciously articu-

lates what qualities contribute to the success of his work and shapes a clearer concept of what is important in the learning process.

The critics of teacher evaluation are many; their chief attack weapon is the plea for:

1. objectivity and fairness in the evaluation of students, and
2. uniformity and consistency of marking standards.

One of the most devastating counter-arguments in recent times has been Sir Alec Clegg's famous remark in his 1966 address "Education: Mind Stock or Fire Kindling?" As a visitor to Canada sponsored by the Canadian Education Association, Sir Alec brought this message to teachers, administrators and parents:

> "The world of education is full of illusions . . . the third is that someone who has not taught the children can more effectively place them in relation to each other than someone who has."

Sir Alec argues that it is necessary to get rid of standardized tests and rely on the teachers' decisions, that there are other methods of promoting comparable standards from school to school than standardized tests, that teachers freed of external controls use "much more varied methods of stimulating learning in their pupils", that intangible qualities such as motivation, curiosity, self-discipline and creativity are far more important than performance on a standardized test in assessing the potential achievement of individual students, and that these can only be assessed by the professional teacher who applies "the knowledge and wisdom which his training has given him". After all, even when a doctor has a personal distaste for a patient, does he allow his personal animosity to overcome his sense of professional ethics or integrity to interfere with the exercise of his skill in diagnosing and treating the patient? Likewise—the teacher.

As far as reasonable consistency in marking standards is concerned, the chief solution lies in:

1. stronger lines of communication
2. more ambitious programs of in-service training
3. distributive evaluation practices, and
4. multiple marking procedures

To be specific:

1. English Department Heads should organize staff meetings for the purpose of discussing marking procedures and standards.

2. Principals should timetable grade level marking periods so that all members of an English staff teaching at a certain level can round table mark a limited number of selected papers, compare standards, and develop consistent procedures.
3. Marking workshops involving teachers of the same grade level from several schools should be organized and held once a term.
4. The O.S.S.T.F. or the Ontario Council of Teachers of English could be asked to publish annually a booklet of student writings with marking annotations and marks assigned by a substantial cross-section of experienced teachers.
5. A major essay assigned during the year to determine part of the student's mark could be handled by the multiple marking process.

The multiple marking process is especially excellent in producing a valid mark. Traditional approaches to essay marking have stressed the methodical consideration of the various aspects of an essay, analytical assessment, the eliminating of differences and the minimizing of individual subjectivity. In sharp contrast, the advocates and researchers of multiple marking have taken "the revolutionary step of acknowledging the value of differences between markers." In Examination Bulletin No. 12, the account of several experiments published by the Schools Council in England and entitled "Multiple Marking of English Composition", the new thinking is explained this way (p. 11):

> "In this sense a multiple mark could in fact be (considered) a composite mark, a consensus to which judges contribute their particular sensibilities."

Multiple marking is based on total impression by a marking team. A research educationalist by the name of Wiseman as early as 1949 demonstrated that "the method of total impression will yield sounder judgements than will analytic methods." (p. 5). Another experiment carried out in England by J. N. Britton in 1960 to compare the analytic marking accuracy by single examiners with the marking accuracy of quick impression markers operating as a team, demonstrated a higher correlation for the experimental team (p. 6). It does after all make sense; all students and teachers of English know that the appeal and value of that complex creation an essay is, to some degree, a matter of taste, personal reaction and experience. To achieve, therefore, highly reliable assessments of students' work, the advantage of basing a mark on several samples of the student's work and team marking procedures is self-evident.

Make no mistake about it: it is important that reasonably high and consistent standards be maintained, and this can only be realized if

Grade XIII teachers everywhere accept the challenge and responsibility of communicating, of participating in well organized in-service programs and workshops designed to increase the individual teacher's awareness of the marking procedures and standards of his colleagues.

The alternative in an increasingly dangerous pattern of progression is—

1. accreditation
2. The ignoring by the universities of the Grade XIII teachers' marks and recommendations,
3. the imposition province-wide as the chief criterion for university entrance of the objective-answer multiple-choice, machine-marked, standardized test.

The form of the standardized English examination developing at present in Ontario is the Ontario Standard English Achievement Test (OSEAT), previously designated the Ontario English Composition Achievement Test (OECAT); the Canada-wide counterpart now under development is the Canadian English Proficiency Achievement Test (CEPAT). Both are "relatives" or derivatives of the American College Board Entrance Examination English Composition Test (ECT) developed by the Educational Testing Service at Princeton, New Jersey.

The validity and effects of the standardized tests developed by the E.T.S. to classify large numbers of high school students for University entrance have long been a controversial topic for debate. Of these tests, the January 5th, 1968 issue of *Time* magazine (p. 40) states:

> "A number of educators now contend that the tests are an imprecise indicator of future success—and colleges are relying on them less and less in picking their freshman classes."

The same article quotes Amherst Admissions Dean Eugene Wilson as saying that the S.A.T. scores "do not guarantee the presence of those human qualities and intellectual abilities we value most.". We might well ask: "Why are we in Canada relying more and more on exactly the same kind of test?"

One of the chief dangers of the College Board ECT is that it may be studied for. As the same issue of *Time* magazine puts it (p. 40):

> "Many schools prep their students on the kind of vocabulary and mathematical skills tested by the exams."

Barron's Educational Publishers of 113 Crossways Park Drive, Woodbury, N.Y., sells a 444 page "pressure-cooker" entitled *How to Prepare for College Entrance Examinations.* Even the Department of Measurement and Evaluation of the Ontario Institute for Studies in Education confesses that preparation for the test can improve student performance.

A short paper issued by OISE and entitled "Evaluation of English: Comments on Recent Criticisms of the Ontario English Composition Achievement Test" states on page 8:

> "Students are unaccustomed to objective type questions and it is therefore necessary for the teacher to carefully review with them the directions and examples found in the Student Handbook."

In subjects that lend themselves to predictable, limited, exact and strictly logical patterns of explication, the use of standardized machine-marked tests may be quite valid. After all, 2 plus 2 does equal 4. But English with its wealth of illusion, subtleties of diction, shades of implication, capacity for ambiguity, richness of metaphor and symbol cannot be adequately assessed by dashes on an IBM card. The reflective, perceptive, imaginative student needs words to explain the reasons for his choice. The gifted writer needs words to display his individual expressional powers.

The drawbacks of the externally-set, machine-marked objective answer test can be summarized as follows:

1. Since objective answer tests can be studied for, if the Universities begin to rely heavily upon OSEAT or CEPAT as the chief criterion for university entrance, the teacher who wishes his students to do their best must devote part of his time and energy to preparing his students for the tests; in other words, curriculum and teaching methodology, to some extent at least, fall under the control of the externally-set examination.
2. The mechanical nature of the OSEAT exam as revealed by the sample questions in the 1968 *OACU Student Handbook* would force the English teacher to stress sentence mechanics at the expense of prose appreciation and creative writing assignments.
3. The very nature of the objective-answer, machine-marked exam, if it becomes the main criterion for University entrance, will discourage innovation and originality in the development of new teaching techniques. As Sir Alec Clegg speaking of the national testing system in Britain so forcefully asserts:

   ". . . if the examinations are taken up in the national system, any original kind of teaching, any real advance that is made, or change in old methods, is absolutely outside the examiner's tether. Any original kind of teaching and real advance in method is accordingly killed at once, under the shadow of this death. The dead hand is bearing on it —it cannot live."

Teachers need the challenge of freedom and responsibility to be vital teachers; the security and indolence of uniformity is depressant. Teachers

need to be *trusted*, and treated as professionals, if real progress is to be made through experimentation and innovation in teaching methods. Through local control of courses and methodology, the English teacher can dislodge rigid patterns of traditional teaching and replace them by the personal, participation of the student in the planning and promotion of his own development.

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY the authors

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."